

DOCUMENT RESUME

ED 479 023

CG 032 515

AUTHOR Neace, William P.; Munoz, Marco A.; Olson-Allen, Susan; Weber, John; Johnson, Knowlton

TITLE Pushing the Boundaries of Education: Evaluating the Impact of Research-Based Social Service Programs with Educational Performance Indicators.

PUB DATE 2003-00-00

NOTE 48p.

PUB TYPE Information Analyses (070) -- Reports - Evaluative (142)

EDRS PRICE EDRS Price MF01/PC02 Plus Postage.

DESCRIPTORS Academic Achievement; *Accountability; *Early Intervention; *High Risk Students; Models; *Partnerships in Education; Prevention; *Program Effectiveness; Program Evaluation; School Community Programs; Violence

ABSTRACT

This paper presents a family of studies associated with a large-scale evaluation conducted to assess the impact of research-based models and promising strategies on non-cognitive and cognitive school indicators. In a partnership between community-based organizations and a large urban public school district, project SHIELD provided services to at-risk students. The theory of educational accountability and violence prevention framed this research project. Data from the selected programs revealed that effects of project SHIELD were generally small, but favorable on attendance, tardiness, and disciplinary measures; however, none of the interventions had an impact on academic performance as measured using GPA. Higher dosage of interventions were generally associated with more positive effects on non-cognitive measures. Implications for practice and future research are discussed. An appendix contains 5 data tables. (Contains 31 references.) (Author)

Running head: Pushing the Boundaries of Education

Pushing the Boundaries of Education: Evaluating the Impact of Research-Based Social
Service Programs with Educational Performance Indicators

William P. Neace
Marco A. Munoz
Susan Olson-Allen
John Weber
Knowlton Johnson

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☐ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL HAS
BEEN GRANTED BY

M. Munoz

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

1

Paper prepared for presentation at the annual meeting of the American Educational Research Association (AERA), San Diego, CA, April 12-16.

Abstract

This paper presents a family of studies associated with a large-scale evaluation conducted to assess the impact of research-based models and promising strategies on non-cognitive and cognitive school indicators. In a partnership between community-based organizations and a large urban public school district, project SHIELD provided services to at-risk students. The theory of educational accountability and violence prevention framed this research project. Data from the selected programs revealed that effects of project SHIELD were generally small, but favorable on attendance, tardiness, and disciplinary measures; however, none of the interventions had an impact on academic performance as measured using GPA. Higher dosage of interventions were generally associated with more positive effects on non-cognitive measures. Implications for practice and future research are discussed.

Keywords: Partnerships in Education; Urban Schools; Accountability; Violence Prevention; High Risk Students; Early Intervention; Program Effectiveness; Program Evaluation.

Pushing the Boundaries of Education: Evaluating the Impact of Research-Based Social Service Programs with Educational Performance Indicators

The study of innovations and change in school systems is not new (Hall & Hord, 2001). What is innovative is a rigorous outcome evaluation of a family of school-based interventions that targets the same problem behavior among a school age population. Further, it is innovative for a school system to evaluate the effectiveness of a system of interventions (referred to as research or science-based interventions or best practices) that has varying degrees of empirical evidence of efficacy in a demonstration project or randomized trial. It is also innovative for a school system to rely on evaluation results to make decisions about the sustainability of interventions and best practices.

The opportunity to launch a large scale evaluation of the effectiveness of a family of research-based, school-based interventions and practices that targets a common problem behavior was presented in a large federal grant from a consortium of agencies (Department of Education, Office of Juvenile Justice and Delinquency Prevention, and the Center for Mental Health Services) as part of a Safe Schools/Healthy Students federal initiative. In October 1999, Project SHIELD (Supporting Healthy Individuals and Environments for Life Development) received nearly \$3,000,000 per year for three years to provide students and schools with enhanced infrastructure and comprehensive prevention and early intervention through education, mental health, and social services that promote healthy childhood development and prevent violence, alcohol, and other drug abuse. These services target the development of social skills and emotional resilience necessary for youth to avoid violent behavior and drug use, along with establishing safe, disciplined, and drug free areas within school environments. From the school system's perspective the most important purpose for conducting the evaluations of Project SHIELD interventions and best practices was to provide results that could be used in

making decisions about which interventions should be sustained using school system funds after the federal funding ended.

The Project SHIELD evaluation strategy used here consists of a process and outcome evaluation of selected research-based interventions and best practices using outcome criteria reflective of accountability measures used by the school district to evaluate all of its school-based interventions. The outcome measures chosen (unexcused absences, unexcused tardiness, in-school disciplinary referrals, number of suspensions, and suspension length, and academic performance) served as measures of student behavior for two reasons.

The Framework of Accountability in Education

Student behaviors and performance are of primary concern to school officials interested in continuing the interventions after Project SHIELD funding has ended, and due to the need to satisfy requirements related to school policies on student behaviors and academic performance. School systems need quantifiable measures of student performance effectiveness in a high-stakes accountability environment. In this environment, educational policy making is based on objective information, and although no single means of data collection is sufficient, the data generated by well-designed program evaluations are crucial to an understanding of project impact.

Policy-makers have to refocus the educational reform efforts in general, and the educational excellence issues in specific, toward results on school-related indicators (Munoz, 2002). School districts must guarantee that programs have a demonstrably positive effect on students' key non-cognitive (e.g., attendance, tardies, suspensions) and cognitive (e.g., reading, math, science, and social studies scores) indicators. Although many of the interventions implemented through Project SHIELD have produced favorable results in previous efficacy evaluations, these evaluations have focused mainly on outcomes related to indirect measures of behavior, such as changes in attitudes, knowledge, and self-reported delinquency (e.g., Cooper, Lutenbacher, &

Faccia, 2000; Mytton, DiGuseppi, Gough, Taylor, & Logan, 2002). School administrators are far more interested in outcome results directly related to student behavioral problems and academic performance. Under the conceptualization of accountability as performance, output educational indicators are used to track and evaluate program effectiveness based on student results. Accountability is generally conceived as a demand to judge school programs by their outputs. Accountability systems have been designed to track the progress of educational reforms. The function of an accountability system in education is to monitor and evaluate the performance of the educational programs (Wholstetter, 1991). Schools are expected to make wise use of public resources not only by efficient cost accounting procedures but also by increasing attendance, decreasing suspensions, and showing positive impacts on academic performance.

From a purely theoretical perspective, Murphy (1988) analysis on the relationship between equity and excellence is relevant in this study. It is this conceptualization that integrates the principles of equity and excellence an important issue for the educational reform efforts in an accountability era. The third-generation conceptualization of equity basically comprehends equity as student opportunity to learn; the first-generation only focused on equity as access (i.e. input) and the second-generation focused on equity as school activities and processes. In this regard, this conceptualization goes beyond the traditional input and process focus of prior educational reform efforts and establishes an interesting link with the school efforts toward quality expressed in terms of student achievement. Significant policy changes have to be framed by the conceptualization of equity as excellence in the accountability educational reform era. In this regard, this conceptualization of equity is highly inter-related to accountability understood as performance. Under the conceptualization of accountability as performance, output educational indicators are used to track and evaluate schools.

The move toward greater accountability in education has been one of the hallmarks since the 1970s in public education (Rich, 1985). Since 1974, Levin has argued that performance accountability is concerned with educational outputs. Levin (1974) defined performance accountability as "a periodic report of the attainments of schools and other educational units" (p. 364). Performance reporting includes such measurement techniques as statewide assessments, school report cards, and performance indicators. The overall objective of a performance accountability system is to provide a standard upon which a school can compare its own progress over time. The end results should (a) stimulate actions to improve education, (b) monitor regulatory compliance for state requirements, and (c) produce rewards as well as sanctions to schools (Kirst, 1990).

Accountability is one of the most important approaches to help the right use of collective funds in public education. Through increased surveillance, it is essential to insist that the scarce tax dollars are held accountable for the products they produce through some valid form of student growth measurement (testing). According to Pinkney (1980), through an effective monitoring system, the general public should insist that all classroom teachers are capable of helping children regardless of their cultural background.

Since the 1980s, education has been "rediscovered" and carefully examined. According to Nelson, Palonsky, and Carlson (1990), the schools needed again a reform. Previous generations of education reformers were concerned with making education available to the children of all classes and races. Instead of availability, the 1980s generation was now forced to consider the quality of school experiences. As Adler (1982) argues, the legal mandate for education cannot be satisfy only by guaranteeing all children access to education. To satisfy the educational responsibilities of a democratic society, public education must demonstrate that each student is provided with adequate levels of knowledge and skills. According to Nelson, Palonsky, and Carlson (1990), educational outcomes cannot longer be measured only in

quantity (e.g., years of schooling and the number of high school diplomas granted). Schools must guarantee that education has a demonstrably positive effect on students.

Schools must show that students benefit from their years of attendance, that increased investment in schooling can be measured in greater ability to read, write, and do mathematics, and that moving up the academic ladder from grade to grade is based on merit rather than on social promotion. (Nelson, Palonsky, & Carlson, 1990, p. 286)

The members of the accountability movement believe that answers to qualitative questions must be based on hard data. Schools need quantifiable measures of student performance and teacher effectiveness if accountability was to be implemented. Intelligent policy decisions should be based on objective information, and although no single means of data collection is sufficient, the data generated by well-designed school indicators are crucial to an understanding of school outcomes. Academic indicators became the scientific basis for making decisions about the quality of schools to produce academic learning by students. Academic indicators became the yardstick against which society charted the progress and shortcomings of education, and became the form in which schools reported the status of public education to public officials and parents. Impressionistic data was not sufficient and anecdotal data was not scientific. The accountability era had entered into public education and was here to stay until today.

Accountability and outcome-based education is a culminating demonstration of learning. In that sense, outcome-based education focuses on identifying and defining specified educational results and teaching toward them (Ellis & Fouts, 1994). Clarity of focus means that all curriculum design, instructional delivery, and assessment is geared to what the students should learn at the end of their time in schools. The expectation is that all students should be able to learn significant curriculum in schools; this means expanding the ways and number of times students get a chance to learn. Clarity of focus, expanded opportunity, high expectations, and

design curriculum backward became the four defining principles of an outcome-based education (Brandt, 1993).

The popularity of performance-based accountability systems have grown since *A Nation at Risk* was published (National Commission on Excellence in Education, 1983), with many states using standardized test scores as the indicator of school quality, building incentive regimes involving payment for results, and threatening state takeovers of poorly performing schools. The problem that arises is that such approaches have consistently simplified student learning and ignored the learning conditions and social context (Darling-Hammond, 1993; Ginsberg & Berry, 1998).

Violence Prevention Perspective

The violence prevention literature also informed the development of the stated objective measures. Currently, there is a call for more use of objective outcome measures (Housman, Pierce, & Briggs, 1996; Webster, 1993). Several studies report improvements in knowledge and attitudes related to use of violence to resolve conflicts (e.g., Aber, Jones, Brown, Chaudry, & Samples, 1998; Hausman, Spivak, & Prothrow-Stith, 1995; Oprinas, Parcel, McAlister, & Frankoweski, 1995), or self-reported aggressive behaviors (e.g., Avery-Leaf, Cascardi, O'Leary, & Cano, 1997; DuRant, Barkin, & Krowchuk, 2001; DuRant, Treiber, Getts, McCloud, Linder, & Woods, 1996).

Fewer studies have examined changes in relevant student behavioral indicators, such as suspension rates (Farrell, Meyer, & White, 2001; Housman et al., 1996), disciplinary referrals (Ferrell et al, 2001; Twemlow, Fonagy, Sacco, Gies, Evans, & Ewbank, 2001) or academic performance (Twemlow et al., 2001). Only one study (Ferral et al., 2001) has examined the link between changes in knowledge and attitudes and relevant behavioral indicators, with the results indicating no mediating effects. Thus, it is becoming increasingly important to examine the effectiveness of violence prevention interventions on student behavior and performance. The

following section presents a brief description of the selected Project SHIELD interventions evaluated in this study.

Project SHIELD Interventions

The selected research-based interventions and best practices under evaluation fall into three categories (Tolan, Guerra, & Kendall, 1995): *universal interventions* (those in which participation is a consequence of class attendance and that target students broadly), *selective interventions* (targeting specific groups within a population that have elevated risk for developing a problem), and (3) *indicated interventions* (targeting high-risk members of a population that are exhibiting detectable signs or symptoms of developing a problem). The universal interventions evaluated here are SMART Moves, and Second Step. The selective interventions are LEEP Mentoring, Big Brothers/Big Sisters Mentoring, and Primary Mental Health. The indicated interventions are Multi-Systemic Therapy and Functional Family Therapy. A brief description of each intervention follows.

Universal Interventions

SMART Moves. SMART (Skills Mastery and Resistance Training) Moves is a research-based, universal program (grades 6, 8, 10, and 12) consisting of a series of validated prevention program components for children aged 10-15 years and their parents. The Salvation Army Boys and Girls Clubs are implementing SMART Moves to increase awareness of Alcohol, Tobacco and Other Drug (ATOD) risks, reduce risky behaviors, and improve academic achievement. The program components focus on alcohol, tobacco, other drugs (Stay SMART), pregnancy, peer pressure, decision-making and life planning skills (Start SMART), conflict resolution and dealing with anger (Street SMART), healthy lifestyle choices for girls (SMART Girls), and parental support (SMART Parents). Implementing these program components at the transition period of elementary through middle school ages provides a critical age-appropriate

support system that addresses the issues students face as they experiment with their approaching adult rights, roles, and responsibilities.

Second Step. Second Step is a research-based, universal violence prevention program for K-middle school-aged children. Seven Counties Services is implementing Second Step in grades 1 to 3 in order to increase children's ability to identify what others are feeling, take others' perspectives, and respond empathically with others. The program also hopes to decrease impulsive, aggressive, and angry behavior. The 28 lessons focus on precursor behaviors that are incompatible with violence: empathy, impulse control, problem solving, and anger management. Lessons are developmentally appropriate in content and delivery with ample opportunity for students to model, practice, and reinforce their prosocial behavior.

Selective Interventions

LEEP Mentoring. The Louisville Education and Employment Partnership (LEEP) is a selective program that uses a mentoring team concept, and has been implemented at participating middle schools (grades 6-8) in the school district. LEEP, including mentoring, has been expanded in two Project SHIELD high schools (grades 9-12) to include an additional 55 students each. Through LEEP activities (including tours, career workshops, job fairs, and breakfast clubs), group mentoring, and one-on-one mentoring, the program hopes to improve school attendance rates, cognitive outcomes, graduation rates, and employment, post-secondary education, or military transition.

Big Brothers/Big Sisters Mentoring. Big Brothers/Big Sisters School-Based Mentoring is a research-based, selective program focusing on case managed mentoring by adult volunteers of 100 elementary students (K-5) each year who are considered high risk to engage in risky behaviors which include violence and substance use. Through its mentoring program, Big Brothers/Big Sisters hopes to increase the children's confidence, competence, caring, and cognitive outcomes, and to decrease behavioral problems. The program (1) recruits, screens,

and trains adult volunteers, (2) interviews and assesses student applicants and their parents, (3) makes and supervises matches between the adults and students, and (4) closes the matches when timelines or situations dictate.

Primary Mental Health. Primary Mental Health Project is a research-based, selective program. This early detection and prevention program for preschool and primary grades is being implemented by JCPS. It is a nationally recognized model out of Rochester, New York that has been replicated in over 200 cities since 1957. Trained child associates (supervised by school psychologists) provide individual interactive play sessions to children (grades 1-2) who begin showing adjustment difficulties. Learning skills and other school-related competencies are enhanced, and social, emotional, and school adjustment problems are reduced.

Indicated Interventions

Multi-Systemic Therapy. Multi-Systemic Therapy (MST) is a nationally recognized, research-based, indicated program for chronic, violent, or substance abusing juvenile offenders aged 12-18 (grades 9-12). MST is being implemented by Seven Counties Services, to improve school attendance, and reduce suspensions, antisocial behavior, and youth criminal activity. Replication studies show extensive improvement in family relations, decreased behavioral problems, reduced substance abuse and a reduced three-year recidivism rate. The program considers the complex network of systems that affect individual, family, and other relationships. Parents and students experience strategic and structural family therapy, behavioral parent training, and cognitive behavior therapies to focus attention and direct change.

Functional Family Therapy. Functional Family Therapy (FFT) is a nationally recognized research-based, indicated program that has been applied successfully in various contexts (e.g., rural, urban, multicultural, international) and treatment systems (e.g., clinics, home-based programs, independent providers). Seven Counties Services is implementing FFT to engage and retain families and targeted youth in prevention/intervention activities, develop long-term

behavior change patterns, and enhance families' abilities to impact multiple systems in which the families are embedded. This program's success has been replicated and demonstrated for over twenty-five years. The program is designed for individual youth ages 11-18. The youth present a wide range of acting-out behaviors and are at risk for, or are presenting behaviors associated with delinquency, violence, substance use, conduct disorder, oppositional-defiant disorder, or disruptive behavior disorder. Therapists administer the program by responding promptly to family expectations, teaching parenting, problem solving and conflict management skills, and initiating clinical and community linkages with the families. The program focuses on enhancing protective factors and reducing risk factors through engagement, motivation, assessment, behavioral change, and generalization of results to apply to the family's needs.

This paper will inform school administrators about the impact of these programs on key performance indicators associated with the school environment. The present study was designed to contribute to existing theoretical and practical knowledge about the effectiveness of a family of research-based and best practices interventions on school-related research issues not previously addressed in the literature. According to Chen and Rossi (1983), the use of theoretical models in program impact assessment can heighten the power of experimental designs and compensate for some deficiencies in the quasi-experimental designs. In the language of Rossi, Freeman, and Lipsey (1999), the distal outcome of student achievement reflects a conceptual hypothesis of the implicit program impact theory of Project SHIELD.

This report presents the final results of a process and outcome evaluation of two cohorts of students who participated in selected Project SHIELD interventions during the course of the 2000-2001 and 2001-2002 school years. The specific research questions that guided the evaluation are outlined below.

Process Questions

There are three process research questions that are specifically relevant to program evaluation. The process questions relate to the analysis of implementation quality as operationally defined by coverage (number targeted vs. number reached), dosage (intensity of the intervention), and implementation fidelity (implemented as designed). These questions are listed as follows:

Question P-1: To what extent does the intervention/best practice reach an adequate sample of the intended targeted population (Coverage)?

Question P-2: What discrepancies (if any) exist between the planned and actual level of exposure to intervention/best practice activities (Dosage)?

Question P-3: What discrepancies (if any) exist between the intervention/best practice design and how the design was implemented (Implementation Fidelity)?

Outcome Questions

The Project SHIELD outcome evaluation addresses the following questions:

Question O-1. Is there a reduction in non-cognitive outcomes (number of unexcused absences, tardies, ISAPs, and suspensions) as a result of participation in Project SHIELD interventions in comparison to a matched control group?

Question O-2. Is there an increase in cognitive outcomes as a result of participation in Project SHIELD interventions in comparison to a matched control group?

Question O-3. Does dosage (length of time a student received services) predict change in non-cognitive and cognitive outcomes?

All programs being evaluated began full operation in the fall semester of 2000. Each of these interventions was evaluated using the methodology described more fully in the next section.

Method

The methodology presented in this section of the study was used to address specific research questions concerning one of Project SHIELD's stated goals: To positively impact the safety and well-being of students. Outcome evaluation is necessary in order to determine whether an intervention has the desired effect on relevant criteria by which an it is to be judged (Rossi, Freeman, & Lipsey, 1999). The outcome evaluation employed in the present study uses a pre-post matched control group design to assess outcomes for participants in the seven interventions being evaluated. Due to the inability to randomly assign individuals to intervention and control groups, a two-level matching procedure was employed in order to add rigor to the internal validity (Cook and Campbell, 1979) of the evaluation. The first level involved the selection of a subset of Jefferson County public schools in which Project SHIELD was not being implemented but which were similar in global characteristics (e.g., percentage of students involved in the free and reduced lunch program, percentage of students in special education programs, percentage of students suspended) to Project SHIELD schools. These schools served as the basis for the second level of matching which took place at the individual level. The control group "partners" for each of the intervention group individuals were selected from the subset of control schools by database software authored by Mr. Mark Shuler of MBS Technologies, Inc. Students were matched on three demographic characteristics considered important by key district-level administrators: (1) Exceptional Childhood Education (ECE) status, (2) gender, and (3) age. This procedure resulted in providing a unique, matched control group for each of two cohorts of students receiving Project SHIELD services within each of the interventions being evaluated. Students in Cohort 1 began receiving services in the fall semester of the 2000-2001 school year. Students in Cohort 2 began receiving services in the fall semester of the 2001-2002 school year. An analysis of the success of the matching procedure is presented in the results section.

Participants

As mentioned above, the matching was implemented using students from matched control schools along three dimensions: ECE status, gender, and age. Table 1 presents the number of SHIELD participants and control students for each of the interventions evaluated, broken out by cohort. Overall, the matching success rate was 96% for Cohort 1 and 97% for Cohort 2. Chi-square analysis was used to test for statistically significant differences between the SHIELD and control groups on categorical variables (ECE status, gender), and a one-way analysis of variance was used to determine whether the two groups of students statistically differed in age. The matching results, presented in Table 2 (on the following page), summarizes the findings. As can be seen, the distribution of students in the control groups was very similar to the distribution of students in the corresponding SHIELD interventions.

Table 1

Number of participants and controls for Project SHIELD interventions

COHORT 1		
Intervention	Number of Control Students	Number of SHIELD Students
SMART Moves	1708	1,814
Second Step	918	922
LEEP Mentoring	102	119
Functional Family Therapy	69	75
Primary Mental Health	233	233
Big Brothers/Big Sisters	151	152
Multi-Systemic Therapy	52	57
TOTAL	3,233	3,372
COHORT 2		
Intervention	Number of Control Students	Number of SHIELD Students
SMART Moves	2329	2416
Second Step	1200	1225
LEEP Mentoring	97	100
Functional Family Therapy	24	25
Primary Mental Health	443	450
Big Brothers/Big Sisters	179	180
Multi-Systemic Therapy	34	36
TOTAL	4306	4432

Chi-square analyses on categorical demographics and analysis of variance on the continuous age measure conducted for each intervention indicated that the control groups did

not significantly differ from the corresponding SHIELD groups on ECE status, gender, or age for either Cohort 1 or Cohort 2. From these analyses, we conclude that the two groups are well-matched on the selected criteria.

Table 2

Results of the computerized matching procedure for each Project SHIELD intervention for Cohort 1 and Cohort 2

COHORT 1						
Intervention	ECE Status			Gender		Age (in years)
	None	Mental / Physical	EBD	Female	Male	
SMART Moves	1230 (68%)	522 (29%)	62 (3%)	835 (46%)	979 (54%)	14.49
Control	1199 (70%)	466 (27%)	43 (3%)	793 (46%)	915 (54%)	14.49
Second Step	733 (80%)	177 (19%)	12 (1%)	448 (49%)	474 (51%)	8.29
Control	733 (80%)	175 (19%)	10 (1%)	446 (49%)	472 (51%)	8.30
LEEP	71 (60%)	46 (39%)	2 (2%)	70 (59%)	49 (41%)	18.14
Control	60 (59%)	40 (39%)	2 (2%)	57 (56%)	45 (44%)	18.06
FFT	46 (61%)	19 (25%)	10 (13%)	41 (55%)	34 (45%)	14.60
Control	45 (65%)	16 (23%)	8 (12%)	39 (57%)	30 (44%)	14.62
PMH	169 (73%)	63 (27%)	1 (0.4%)	94 (40%)	139 (60%)	8.82
Control	169 (73%)	63 (27%)	1 (0.4%)	94 (40%)	139 (60%)	8.93
BB/BS	109 (72%)	38 (25%)	5 (3%)	68 (45%)	84 (55%)	10.98
Control	109 (72%)	38 (25%)	4 (3%)	68 (45%)	83 (55%)	10.99
MST	25 (44%)	23 (40%)	9 (16%)	18 (32%)	39 (68%)	17.43
Control	24 (46%)	21 (40%)	7 (14%)	18 (35%)	34 (65%)	17.57
COHORT 2						
Intervention	ECE Status			Gender		Age (in years)
	None	Mental / Physical	EBD	Female	Male	
SMART Moves	1681 (69.6)	632 (26.2)	103 (4.3)	1102 (45.6)	1314 (54.4)	15.6
Control	1648 (70.8)	595 (25.5)	86 (3.7)	1067 (45.8)	1262 (54.2)	15.6
Second Step	930 (75.9)	266 (21.7)	29 (2.4)	592 (48.4)	632 (51.6)	9.2
Control	915 (76.3)	259 (21.6)	26 (2.2)	580 (48.3)	620 (51.7)	9.2
LEEP	55 (55)	41 (41)	4 (4)	59 (59)	41 (41)	19.0
Control	55 (56.7)	38 (39.2)	4 (4.1)	56 (57.7)	41 (42.3)	18.8
FFT	13 (52)	10 (40)	2 (8)	11 (44)	14 (56)	15.6
Control	13 (54.2)	10 (41.7)	1 (4.2)	11 (45.8)	13 (54.2)	15.5

Table continues...

Table 2 continued

Results of the computerized matching procedure for each Project SHIELD intervention for Cohort 1 and Cohort 2

Intervention	COHORT 2 ECE Status			Gender		Age (in years)
	None	Mental / Physical	EBD	Female	Male	
PMH	310 (68.9)	130 (28.9)	10 (2.2)	197 (43.8)	253 (56.2)	9.4
Control	306 (69.1)	128 (28.9)	9 (2)	193 (43.6)	250 (56.4)	9.3
BB/BS	122 (67.8)	50 (27.8)	8 (4.4)	82 (45.6)	98 (54.4)	11.9
Control	122 (68.2)	49 (27.4)	8 (4.5)	82 (45.8)	97 (54.2)	11.7
MST	15 (41.7)	14 (38.9)	7 (19.4)	15 (41.7)	21 (58.3)	17.6
Control	15 (44.1)	12 (35.3)	7 (20.6)	15 (44.1)	19 (55.9)	17.5

Instrumentation

Process Data. The process measures of coverage, dosage, and fidelity were used to evaluate the implementation quality of each SHIELD intervention. Coverage is defined as the reach, or number of units (e.g., students, parents, schools), that received each service offered by each intervention. These indicators varied with each intervention. Dosage is defined as the number of hours of service delivered and received by participants of each intervention. Finally, implementation fidelity provides a measure of the extent to which intervention activities were implemented according to the SHIELD plan and/or as they were designed by the developer. One process instrument was developed for each of the interventions in this study, tailored to process measures relevant to each intervention. Data (coverage, dosage and implementation fidelity) were collected for all three years via a self-administered questionnaire completed by each intervention director. Data were collected at the end of each project year.

Outcome Data. The outcome evaluation relied on secondary data collected by school personnel and maintained in the school systems' Management Information System using Teradata database software. Detailed student information on unexcused absences and unexcused

tardiness is collected on a daily basis by the school system, and maintained in the Teradata warehouse. The Teradata warehouse can also be used to track a number of more direct indicators of problematic behavior. In-school alternative program (ISAP) referrals (a form of in-school suspension) are tracked on a daily basis. Generally speaking, the number of days a student spends in this form of punishment could serve as a rough indicator of his or her behavioral problems. Two other measures of behavioral problems are also available. The first measure, number of out-of-school suspensions, serves as an indicator of the extent of problem behaviors exhibited by the student. The more total number of suspensions per semester, the more problematic the behavior of the student. The second measure extracted is the total number of days suspended, which serves as an indicator of the seriousness of the offense that resulted in a suspension. Generally speaking, the more serious the offense, the more days the student is suspended. Infractions for which a suspension could result range in seriousness from failure to sign out of school to terroristic threatening (e.g., bomb threat directed at the school). Finally, the Teradata warehouse also contains some academic information for middle and high school students that can be used as an indicator of academic performance. These data include academic grades in four core content courses (English, Social Studies, Mathematics, and Science), coded on an ordinal scale ranging from F to A+. Grades were recoded into a numeric scale (ranging from 0 to 4.5, representing F to A+, respectively), and used to form GPA (grade point average) by taking the mean of the four core content courses within each semester.

Outcome measures were collected the semester before the students' involvement with Project SHIELD (baseline), after one semester of involvement in the particular Project SHIELD intervention being evaluated (post-intervention short term effects), and after up to four semesters of involvement (post-intervention follow up assessing long term effects). Baseline for Cohort 1 was the spring semester of the 1999-2000 school year, with short term effects assessed in the spring semester of the 2000-2001 school year and long term effects assessed

at the end of the spring semester of the 2001-2002 school year. Baseline for Cohort 2 was the spring semester of the 2000-2001 school year, with assessment of short term effects occurring in the spring semester of the 2001-2002 school year and long term effects assessed in the spring semester of the 2002-2003 school year.

In addition to the outcome measures just described, demographic characteristics of the students were also collected. These included Exceptional Childhood Education (ECE) Status (a general indicator of the student's cognitive, psycho-social and physical functioning), gender, and age (chronological age in years). These measures were collected from both cohorts in order to examine the degree to which students in the intervention groups were matched to control students. Data on free-and-reduced lunch status (in which students are classified according to their ability to pay for their own school meals, ranging from those who receive free meals to those who pay full price for meals), and zip code (area of residence) were also collected to be used as covariates in the analysis. In addition, a dosage measure consisting of the number of weeks a student received Project SHIELD services was collected for Cohort 2 (dosage measures were not available for Cohort 1 students).

Analysis Strategy

Process Analysis. The process data analysis consisted of a comparison of the anticipated results for each intervention activity to the actual results achieved at the end of each semester. The actual results (e.g., fully achieved, partial achieved, not achieved) were determined by using questions for each coverage, dosage, and fidelity measure. In addition, an index was constructed for each of the categories of implementation fidelity (coverage, dosage, and implementation fidelity) by summing the item scores within that category. Descriptive statistics and frequencies were used to

determine the average extent of coverage, dosage, and implementation fidelity for each intervention across the three years of Project SHIELD implementation.

Outcome Analysis. As indicated earlier, the outcome evaluation is an assessment of secondary data maintained in the Teradata warehouse. The data were first examined to test the statistical assumptions (e.g., distributional assumptions of the outcomes, homogeneity of variance, examination of outliers) of the planned analysis procedures (Tabachnik & Fidell, 1996). Since the data were found to be amenable to general linear modeling, the interventions were evaluated using a repeated measures analysis of covariance (ANCOVA) with group (intervention group vs. control group) as a between-subjects factor, and repeated measures on outcome as the within-subject factor. Students were matched on ECE status, gender, and age (see results below); however, pre-existing group differences on eligibility for free or reduced lunch prices (in Cohort 1) and area of residence (in both Cohort 1 and Cohort 2) had to be statistically controlled by entering them as covariates in the models. Covariates were used here principally to rule out the effects of potentially confounding factors, such as differences in participation in the free-and-reduced lunch program (our proxy variable for socio-economic status), on assessing the change in behavior and cognitive outcomes. For Cohort 1, short-term outcome change was defined as the difference in outcome measures from the Spring 2000 semester to the Spring 2001 semester, reflecting short term outcomes. Long-term outcome change was defined as a change in outcome measure from the Spring 2000 semester to the Spring 2002 semester. For Cohort 2, short term change was defined as the difference in outcomes measures from the spring 2001 semester to the spring 2002 semester, and long term outcome change was defined as the differences in outcome measures between the spring 2001 semester and the spring 2003 semester. Separate ANCOVAs were conducted for each of the interventions of interest within each of the cohorts in the study.

A second set of analyses were conducted for Cohort 2. Using service start and end dates, we calculated the number of weeks a student was involved in the intervention and took this result as a measure of dosage. A series of regressions were then conducted within the SHIELD group to assess the relationship between the number of weeks receiving services and the change in the outcome measures. Changes in outcomes were calculated by subtracting pre-intervention measures from the post-intervention measures to assess both short term and long term effects.

Results

Process Results

On average across the three years of implementation, 100% coverage was obtained by Second Step, LEEP, Primary Mental Health, and Big Brothers/Big Sisters. SMART Moves achieved 90 percent coverage. Functional Family Therapy and Multi-Systemic Therapy achieved 79% and 77% coverage, respectively. Among the reasons cited for the lower coverage in the indicated interventions were that more families and students were targeted each year, and that intervention referrals are typically low at the beginning of the school year when year 3 data were collected.

Full dosage (100%) was achieved for SMART Moves, LEEP, Primary Mental Health, Big Brothers/Big Sisters, and MST on average across the three years of Project SHIELD. Functional Family Therapy achieved 81% dosage, while Second Step dosage was only 67%. The intervention directors indicated that sessions were still in progress when data were collected, which accounts for the lowered dosage for Functional Family Therapy and Second Step.

On average, implementation fidelity was high. Both Functional Family Therapy and Primary Mental Health achieved 100% implementation fidelity. Multi-Systemic Therapy (98%) Big Brothers/Big Sisters (92%), SMART Moves (87%), and LEEP (86%) were also implemented

with a high degree of fidelity. Implementation fidelity was somewhat lower for Second Step (74%); however, the implementation director indicated that the intervention was adapted so as not to duplicate other services currently provided in Project SHIELD schools.

In summary, with few exceptions, each intervention reached its intended population, was fully implemented, and was implemented with high fidelity. This implementation quality was judged to be adequate.

Outcome Results

The results of the main outcome analyses are reported below for each of the Project SHIELD interventions separately. Short term results (one semester) are reported first for each cohort within each intervention, followed by the long term results (up to four semesters). The results tables are organized as follows: For Table 3 (short term results), column 1 lists the outcomes. Columns 2 and 3 show the means for each outcome in the pre-intervention semester (Spring 2000 for Cohort 1 and Spring 2001 for Cohort 2) and the post intervention semester (Spring 2001 for cohort 1 and Spring 2002 for Cohort 2) for the intervention groups. Columns 4 and 5 show these means for the control groups. The F-test value for the group by time interaction is shown in column 6, and column 7 presents the effect size. Effect sizes, a measure of the amount of pre-post change attributable to a factor, are discussed in text where results are statistically significant. According to Cohen (1988), small effect sizes are a bivariate correlation of .10 or less, medium effect sizes are a bivariate correlation of about .30, and large effect sizes are a bivariate correlation of .50 or greater. Table 4 (long term results) outcomes are organized similarly, except that the means for the follow-up assessments are presented instead of the post-intervention means in columns 2 and 5 for intervention groups and controls, respectively.

Short Term Change in Non-Cognitive and Cognitive Outcomes

Table 3 displayed in Appendix A presents the short term outcome results for each cohort for the seven interventions evaluated. For each intervention, Cohort 1 results are presented first, followed immediately by Cohort 2 results. Each intervention will be discussed in turn.

SMART Moves.

- Cohort 1: Differential change in behaviors between students in SMART Moves and controls was significant but in the wrong direction for four of the five non-cognitive outcome measures. For students in SMART Moves, unexcused tardiness, number of days suspended, and total number of suspensions increased more for the intervention group than for the control group. Effect sizes for these differences were small, ranging from .07 for the change in suspensions to .13 for the change in tardiness. ISAPs increased more for control students than for students in the first cohort of SMART Moves (effect size = .05). Change in cognitive outcomes (GPA) did not differ between intervention and control students in Cohort 1.
- Cohort 2: More favorable results are evident in Cohort 2. Unexcused absences and unexcused tardiness decrease for students in the second SMART Moves cohort, relative to controls. Students in SMART Moves also exhibited a decrease in number of suspensions and number of days suspended, relative to the increase observed in the control group. Effect sizes for these findings were small, ranging from .04 (for the decrease in unexcused absences) to .12 (for the decrease in suspensions). GPA decreased slightly for students in SMART Moves relative to controls; however, the effect size was very small (.04).

Second Step.

- Cohort 1: Students in the first Second Step cohort exhibited a greater decrease in unexcused absences, relative to the control group. The effect size was small (.13). No other group differences in Cohort 1 were significant.
- Cohort 2: Students in Second Step showed a decrease in the number of unexcused absences relative to controls (effect size = .05), and also exhibited a decrease in unexcused tardiness relative to an increase seen in the control group (effect size = .06). No other group differences were statistically significant.

LEEP.

- Cohort 1: Students in LEEP Cohort 1 exhibited a decrease in unexcused tardiness compared to controls, who exhibited an increase (effect size = .17).
- Cohort 2: Students in LEEP Cohort 2 showed a decrease in unexcused absences, relative to controls that approached significance (effect size = .12). No other group differences for Cohort 2 were significant.

Functional Family Therapy.

- Cohort 1: Students in Functional Family Therapy exhibited an increase in unexcused absences (effect size = .21), and an increase in unexcused tardiness (effect size = .19), relative to controls. The Cohort 1 intervention group also exhibited an increase in number of suspensions (effect size = .18), relative to the observed decrease for control students.
- Cohort 2: There were no statistically significant effects noted between Functional Family Therapy Cohort 2 students and their control group.

Primary Mental Health.

- Cohort 1: There were no statistically significant group differences noted in the short-term analysis of Cohort 1 students in Primary Mental Health.

- Cohort 2: Students in Primary Mental Health Cohort 2 exhibited a decrease in unexcused tardiness relative to an observed increase in the control group (effect size = .09).

Big Brothers/Big Sisters Mentoring.

- Cohort 1: Students in Big Brothers/Big Sisters Mentoring exhibited a decrease in unexcused tardiness, while control students exhibit a slight increase (effect size = .17).
- Cohort 2: Students in Big Brothers/Big Sisters Mentoring did not significantly differ from controls students on any of the outcomes assessed in the short term.

Multi-Systemic Therapy.

- Cohort 1: Students in Multi-Systemic Therapy exhibited a reduction in unexcused tardiness relative to matched controls, who exhibited an increase (effect size = .23). No other group differences were significant for students in Cohort 1.
- Cohort 2: There were no significant group differences noted in the short term outcome analysis.

Summary of Short Term Results

- In the short term, some of the Project SHIELD interventions appeared to positively impact unexcused absences and unexcused tardiness. Relative to matched control students, students in SMART Moves (Cohort 2), Second Step (Cohorts 1 and 2), and LEEP Mentoring (Cohort 2) exhibited decreases in the number of unexcused absences. Decreases in unexcused tardiness were observed for students in SMART Moves (Cohort 2), Second Step (Cohort 2), LEEP (Cohort 1), Primary Mental Health (Cohort 2), Big Brothers/Big Sisters (Cohort 1), and Multi-Systemic Therapy (Cohort 1), relative to matched controls.

- Among findings that were not positive were increases in unexcused absences, unexcused tardiness, suspensions, and number of days suspended for students in SMART Moves (Cohort 1) relative to controls. Students in Functional Family Therapy (Cohort 1) also exhibited increased unexcused absences relative to control students. Effect sizes for these findings were small, however.

Long Term Change in Non-Cognitive and Cognitive Outcomes

Table 4 in Appendix A presents the long term outcome results of the outcome evaluation. As in the preceding discussion, Cohort 1 results are presented first, followed by Cohort 2 results for each intervention separately.

SMART Moves.

- Cohort 1: Unexcused absences increased for both SMART Moves and control students, but the magnitude of increase was larger for students in SMART Moves than for control students (effect size = .05). Analysis of Long term changes in non-cognitive measures for Cohort 1 indicated that ISAPs increased more for control students than for students in SMART Moves (effect size = .04). This is a very small positive effect.
- Cohort 2: Analysis of long term change in SMART Moves showed that SMART Moves students decreased in unexcused absences (effect size = .06), unexcused tardiness (effect size = .10), number of suspensions (effect size = .12), and number of days suspended (effect size = .10) relative to increases in these outcomes observed in the control group. GPA for students in SMART Moves decreased sharply compared to controls (effect size = .39).

Second Step.

- Cohort 1: In the long-term for Cohort 1, unexcused absences decreased for students in Second Step but increased for control students (effect size = .14). No other group differences were significant.
- Cohort 2: Students in Second Step showed a decrease in unexcused absences (effect size = .04) and unexcused tardiness (effect size = .07) relative to controls. Students in Second Step showed no change in the number of ISAPs, which increased for control students (effect size = .04). This is a very small positive effect.

LEEP Mentoring.

- Cohort 1: Students exhibited a decrease in ISAPs, relative to the increase seen in the control group (effect size = .15). Students showed a decrease in unexcused absences, relative to controls that approached significance (effect size = .13). No other group differences were significant in the analysis of long term effects for LEEP Cohort 1.
- Cohort 2: Analysis of long term effects indicated that students in LEEP exhibited a decrease in unexcused absence relative to controls (effect size = .18).

Functional Family Therapy.

- Cohort 1: Students in Functional Family Therapy Cohort 1 exhibited a slight increase in ISAPs relative to control students, who exhibited a substantial increase (effect size = .20). This is a small to medium size positive effect.
- Cohort 2: No group differences were statistically significant.

Primary Mental Health.

- Cohort 1: As in the analysis of short-term effects, there were no statistically significant group differences noted in the long-term analysis of the effects of Primary Mental Health Cohort 1.

- Cohort 2: Students in Primary Mental Health exhibited a decrease in unexcused tardiness while controls exhibited an increase (effect size = .05).

Big Brothers/Big Sisters Mentoring.

- Cohort 1: There were no significant group differences in the analysis of long term effects.
- Cohort 2: Same results for Cohort 1.

Multi-Systemic Therapy.

- Cohort 1: Significant Cohort 1 group differences in the change in unexcused absences (effect size = .33), number of days suspended (effect size = .25), and total number of suspensions (effect size = .21) were found. Relative to controls, students in MST exhibited medium size positive effects on each of these outcomes.
- Cohort 2: Similar results are noted for Cohort 2. Relative to controls, students in Multi-Systemic Therapy exhibited decreases in ISAPs (effect size = .21) and number of suspensions (effect size = .22).

Summary of Long Term Results

- Analysis of long term results indicated that Project SHIELD had an impact across the entire range of non-cognitive measures. Compared to controls students, students in SMART Moves (Cohort 2), Second Step (Cohorts 1 and 2), LEEP (Cohort 2), and Multi-Systemic Therapy (Cohort 1) exhibited positive effects on unexcused absences. Decreases in unexcused tardiness were seen for students in SMART Moves (Cohort 2), Second Step (Cohort 2), and Primary Mental Health (Cohort 2), compared to control students. Change in ISAPs also show positive intervention effects for students in LEEP (Cohort 1), Functional Family Therapy (Cohort 1), and Multi-Systemic Therapy (Cohort 2). Suspensions decreased for students in SMART Moves (Cohort 2), and Multi-Systemic Therapy (Cohorts

1 and 2). Finally, number of days suspended decreased for students in SMART Moves (Cohort 2) and Multi-Systemic Therapy (Cohort 2) compared to controls.

- For students in Second Step (Cohort 2), little change in suspensions was observed in comparison to an increase in suspensions for control students. The only noted change in cognitive outcomes was that students in SMART Moves (Cohort 2) exhibited a medium size decrease in GPA relative to the control group.

Effect of Dosage on Short Term and Long Term Outcomes (Cohort 2)

Table 5 presents the results of dosage on short term and long term outcomes for students in Cohort 2. Recall that dosage is the number of weeks a student participated in Project SHIELD, calculated from intervention start and end dates. Table 5 is organized as follows: Column 1 shows the intervention and the outcome measure for which an effect was found. Column 2 contains the regression coefficients for the change in outcome measure and for dosage in the short term analysis. Column 3 contains the standardized regression coefficient and reflects the effect size for the dosage effect on the short term behavioral change. Columns 4 and 5 contain the regression coefficients and standardized regression coefficients for the long term analysis of behavioral change. Prior to conducting the regression analysis, change scores on each outcome were calculated by subtracting pre-intervention levels from post-intervention levels (constituting short term change) and pre-intervention levels from follow-up levels of the behaviors (constituting long term change). Change scores (interpreted as changes in the outcomes) were then regressed on dosage to produce the regression coefficients presented in the table.

Two pieces of information are needed in order to interpret the results. The first piece of information is direction of change in the non-cognitive outcomes (a favorable change reflects a decrease in the behavior being studied and an undesirable change indicates an increase). The second piece of information needed is the sign of the regression coefficient. A negative sign for

dosage indicates an inverse relationship between the change in behavioral outcome and number of weeks students received services, while a positive sign for dosage indicates that it had a direct relationship to change in behavioral outcome. Thus, there are two cases that indicate favorable dosage effects. In Case 1, the behavior decreases and dosage is positive (indicating that *more* dosage is related to *greater* decreases in the behavior). In Case 2, the behavior increases and dosage is negative (indicating that *less* dosage is related to a *greater* increase in the behavior). With this in mind, the results presented in Table 5 can be meaningfully addressed. Only significant results are shown.

Big Brothers/Big Sisters: The results indicate that increased dosage favorably impacted ISAPs for students in Big Brothers/Big Sisters in the long term analysis (effect size = $-.15$). While there was an overall increase in ISAPs, decreased dosage was related to greater increases in ISAPs for students in Big Brothers/Big Sisters.

LEEP Mentoring: Overall decreases in unexcused absences and unexcused tardiness are evident in both short term and long term analyses. Increased dosage was associated with greater decreases in unexcused absences in the short term, which is favorable effect (effect size = $.40$) and the long term (effect size = $.42$). Increased dosage was also associated with decreased unexcused tardiness in the short term (effect size = $.33$) and the long term (effect size = $.37$). While number of suspensions and number of days suspended increased in short term, decreased dosage was associated with increased suspensions (effect size = $-.26$) and increased number of days suspended (effect size = $-.29$). Thus, dosage favorably impacted change in the number of suspensions and the number of days suspended for students in LEEP Mentoring.

Multi-Systemic Therapy: Both number of suspensions and number of days suspended decreased in the short term; however, increased dosage was associated with larger decreases in suspensions (effect size = $.35$) and days suspended (effect size = $.41$). Thus, students with

more dosage exhibit greater decreases in number of suspensions and number of days suspended than do students with lesser dosage in the short term analyses.

Second Step: There was an overall decrease in unexcused tardiness for students in Second Step in both the short and long term analyses. Increased dosage was associated with greater decreases in unexcused tardiness in the short term (effect size = .06) and the long term (effect size = .08).

Summary of Dosage Results

The results indicate that dosage favorably impacted ISAPs for students in Big Brothers/Big Sisters in the long term analysis. Dosage was also associated with reductions in unexcused absences (short and long term analyses), unexcused tardiness, suspensions and number of days suspended (in the short term analysis) for students in LEEP Mentoring. Dosage was associated with reductions in number of suspensions and days suspended for students in Multi-Systemic Therapy in the short term analysis. For students in Second Step, dosage was associated with reduced unexcused tardiness in both the short term and the long term analyses.

Discussion

In October 1999, Project SHIELD (Supporting Healthy Individuals and Environments for Life Development) received nearly \$3,000,000 from a consortium of federal agencies (Department of Education, Office of Juvenile Justice and Delinquency Prevention, and Center for Mental Health Services) as part of a Safe Schools/Healthy Students Federal Initiative. The award will provide three years of funding (nearly \$9,000,000) to the school district. Project SHIELD aims to provide students and schools with enhanced infrastructure and comprehensive prevention and early intervention, through education, mental health, and social services that promote healthy childhood development and prevent violence, alcohol and other drug abuse. These services target the development of social skills and emotional resilience necessary for youth to avoid

violent behavior and drug use, along with establishing safe, disciplined, and drug free areas within school environments. To that end, Project SHIELD is implementing 17 research-based violence/substance abuse prevention interventions and best practices. A process and outcome evaluation was conducted for a selected set of research-based and best practices interventions.

Three classes of interventions were evaluated: universal interventions, in which participation is a consequence of class attendance, selected interventions, in which participation is based upon at-risk status, and indicated interventions, which target students already exhibiting behavioral and academic problems. The outcome study used a pre-post matched control group design to analyze secondary data obtained from the Teradata warehouse maintained by the school district. An innovative matching procedure was used to individually match control students to SHIELD students on important demographic characteristics (ECE status, gender, and age). Such matching lends a level of rigor to the evaluation of intervention effects that is seldom seen in such research, and increases confidence in the internal validity of the results (Cook & Campbell, 1979). The evaluation examined both short-term and long-term intervention effects for two separate cohorts of students.

Process results indicated that all of the interventions were implemented with a high degree of quality. Outcome results provided some evidence for the effectiveness of Project SHIELD interventions on school-relevant student behaviors and performance. In the short term, some of the Project SHIELD interventions appeared to positively impact unexcused absences and unexcused tardiness. Relative to matched control students, students in SMART Moves (Cohort 2), Second Step (Cohorts 1 and 2), and LEEP Mentoring (Cohort 2) exhibited decreases in the number of unexcused absences. Decreases in unexcused tardiness were observed for students in SMART Moves (Cohort 2), Second Step (Cohort 2), LEEP (Cohort 1), Primary Mental Health (Cohort 2), Big Brothers/Big Sisters (Cohort 1), and Multi-Systemic Therapy (Cohort 1), relative to matched controls. Short term findings that were not positive were

increases in unexcused absences, unexcused tardiness, suspensions, and number of days suspended for students in SMART Moves (Cohort 1) relative to controls. Students in FFT (Cohort 1) also exhibited increased unexcused absences relative to control students. Effect sizes for these findings were small, however.

Analysis of long term results indicated that Project SHIELD had an impact across the entire range of non-cognitive measures. Compared to controls students, students in SMART Moves (Cohort 2), Second Step (Cohorts 1 and 2), LEEP (Cohort 2), and MST (Cohort 1) exhibited decreases in unexcused absences. There was also an larger increase in unexcused absences in the control group than for student in SMART Moves (Cohort 1). Decreases in unexcused tardiness were seen for students in SMART Moves (Cohort 2), Second Step (Cohort 2), and PMH (Cohort 2), compared to control students. Change in ISAPs also show positive intervention effects for students in LEEP (Cohort 1), FFT (Cohort 1), and MST (Cohort 2). Suspensions decreased for students in SMART Moves (Cohort 2), and MST (Cohorts 1 and 2). For students in Second Step (Cohort 2), little change in suspensions was observed in comparison to an increase in suspensions for control students. Finally, number of days suspended decreased for students in SMART Moves (Cohort 2) and MST (Cohort 2) compared to controls. The only noted change in cognitive outcomes was that students in SMART Moves (Cohort 2) exhibited a medium size decrease in GPA relative to the control group. No other significant effects were found for any of the other interventions for which academic measures were available.

The results also indicated that increased dosage positively impacted non-cognitive outcomes in Cohort 2. ISAPs decreased more for students with higher dosage in Big Brothers/Big Sisters. Positive dosage effects were also found for students in LEEP for changes in unexcused absences, unexcused tardiness, suspensions, and number of days suspended. Dosage also positively impacted suspensions and days suspended for students in Multi-

Systemic Therapy. For students in Second Step, positive effects of dosage on change in unexcused tardiness are noted.

As in any research, there are limitations associated with this evaluation study. This study was not conducted as a randomized, controlled trial, which limits the ability to reach firm causal conclusions about intervention effectiveness. Nevertheless, we partially compensated for this limitation by employing a rigorous computerized matching procedure to create individual student controls who were matched to intervention students on such characteristics as learning disability, gender, and age. Thus, the study has strong internal validity, and differential group change may be attributed more to the interventions than to potentially confounding intervening factors (such as differential maturity, history, and the like discussed by Cook and Campbell, 1979). When random assignment is not possible, future evaluation studies of large comprehensive intervention projects should use similar procedures to control for extraneous factors in the absence of random assignment.

A second limitation of the study is the amount of missing academic data for students. Initially, the investigation planned to use standardized test scores in order to evaluate intervention effects on student academic performance. As the evaluation unfolded, however, it came to our attention that standardized tests are not administered in a manner that would allow for an examination of change in performance at the individual student level. For example, standardized tests in mathematics, reading, and language arts are given to students in grades 3, 6, and 9. The following year, a different standardized test is given to students in grades 6, 7, and 10 for reading and science; however, the reading test is not comparable to the reading test taken the prior year. Students in grades 2 and 3 are given standardized mathematics and reading tests for diagnostic rather than achievement purposes. Thus, standardized test scores could not be used to measure change in academic proficiency for students longitudinally. Semester grades in core content areas (mathematics, English, social studies, and science)

could serve as measures of change in academic performance over time, however, and were thus chosen as the next best indicator. Nevertheless, necessary academic information for many of the students in both intervention and control groups were not available in the Teradata warehouse. The percentage of missing academic data was moderately high (over 20 percent in some cases). Loss of data affected both statistical power as well as limiting generalizability to the rest of the sample. In addition, elementary school students are not given grades in the traditional sense. As a result, student performance measures (grades) for elementary schools are not maintained in the Teradata warehouse. As a result of the amount of missing academic data, changes in student performance could not be adequately measured.

A third limitation is that the amount of individual dosage to a particular intervention could not be gauged in Cohort 1; thus, the average change in outcomes within the intervention groups represents changes for students who have various levels of exposure to the intervention. It is not possible to tease apart students who completed the intervention in question from students who may have participated for only a short period of time for Cohort 1. As a result, changes in outcomes may have been attenuated by varying degrees of exposure to the intervention being evaluated. This limitation is more reflective of the evaluation of the selective and indicated interventions than of the universal interventions, however. Part of this limitation was addressed in the analysis of Cohort 2 since data allowed for calculation of a dosage measure. It would be of benefit to future evaluations of large-scale intervention projects to ensure that some mechanism is in place to assist investigators in tracking individuals exposed to interventions in order to examine differential effects of dosage on outcome. In addition, assessment of the differences in outcomes for students who successfully complete an intervention versus those who do not successfully complete an intervention should also be evaluated. In the present study, we were unavailable to study the effects of successful completion due to exclusion of students who failed to complete an intervention from the sample.

Finally, as noted previously, the present study was designed to examine the effect of violence prevention interventions on indicators of global behavioral and performance measures, rather than to examine outcomes specifically targeted by the interventions (e.g., social skill development, emotional resiliency, conflict resolution skills, knowledge of consequences of violence, and norms related to use of violence). As a result, theoretically relevant intervention effects have been overlooked in favor of examining school-relevant outcomes. Future studies should include both types of outcome measure in order to more fully examine intervention effects.

In conclusion, this study suggests that the Project SHIELD interventions had small but favorable impacts on non-cognitive outcomes relevant to school policy—attendance, tardiness, and disciplinary referrals; however, there is no evidence that these interventions positively impacted the cognitive outcome (GPA). Results also suggest that the SMART Moves, Second Step, and Multi-Systemic Therapy had more favorable results on non-cognitive outcomes than did the other interventions evaluated. Results associated with the universal interventions also appeared to be more consistent across short and long term assessments than those associated with the selective and indicated interventions. These interventions may have beneficial effects on students and their families in ways that were not assessed in the present study, however. Taken together, long term results were generally better across the interventions than short term results, and this was especially true for Cohort 2. The results also suggest that higher dosage was associated with favorable changes in non-cognitive outcomes for Cohort 2, and these changes were particularly evident for students in LEEP mentoring. A majority of the interventions could have increased their favorable effects by increasing dosage.

There was consistent evidence as to the effects of these four interventions on policy-relevant outcomes important to administrators in a high stakes accountability environment where objective, quantifiable results of interventions on student behavior and performance are needed. Therefore it

is recommended that Smart Moves, Second Step, LEEP, and Multi-Systemic Therapy be expanded to other JCPS schools deemed at-risk (e.g., SHIELD control schools) and who are not presently receiving these services. Further, expanding these interventions could be facilitated by exploring grant opportunities to provide the necessary funds to implement these interventions on a larger scale. For the remaining evidence-based interventions and best practices, it is recommended that additional evaluation results be produced to assess their impact on violence and/or substance use before deciding whether or not expand them to other JCPS schools.

It is also recommended that special attention be given to the intensity of the intervention (i.e., dosage) in that higher dosages appeared to yield more favorable outcomes across some of the Project SHIELD interventions we evaluated. In addition, for the future we recommend that careful and continued evaluation of the implementation of the more successful interventions be conducted in order to fully document their impact on student behavior and academic success.

Finally, it is recommended that the scope of the outcome evaluation be expanded to include changes in knowledge and attitudes about violence and substance abuse among youth, as well as changes in family functioning that are specifically targeted by the selective and indicated interventions. Such changes were beyond the scope of the present study but are nonetheless important to include in order to capture the broader impact of violence and substance abuse prevention programming on youth and their families.

References

- Aber, J. L., Jones, S. M., Brown, J. L., Chaudry, N., & Samples, F. (1998). Resolving conflict creatively: Evaluating the developmental effects of a school-based violence prevention program in neighborhood and classroom context. *Development and Psychopathology*, 10, 187-213.
- Adler, M. J. (1982). *The Paideia proposal: An educational manifesto*. New York: Macmillan.
- Avery-Leaf, S., Cascardi, M., O'Leary, K. D., & Cano, A. (1997). Efficacy of a dating violence prevention program on attitudes justifying aggression. *Journal of Adolescent Health*, 21, 11-17.
- Brandt, R. (1993). On outcome-based education: A conversation with Bill Spady. *Educational Leadership*, 50, 66-70.
- Chen, H. T. & Rossi, P. H. (1983). Evaluating with sense: A theory-driven approach. *Evaluation Review*, 7(3), 283-302.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.
- Cooper, W. O., Lutenbacher, M., & Faccia, K. (2000). Components of effective youth violence prevention programs for 7- to 14-year olds. *Archives of Pediatric and Adolescent Medicine*, 154, 1134-1139.
- Darling-Hammond, L. (1993). *Standards of practice for learner-centered schools*. NY: National Center for Restructuring Education, Schools, and Teaching.
- DuRant, R. H., Barkin, S., & Crowchuk, D. P. (2001). Evaluation of a peaceful conflict resolution and violence prevention curriculum for sixth-grade students. *Journal of Adolescent Health*, 28, 386-393.

- DuRant, R. H., Treiber, F., Getts, A., McCloud, K., Linder, C. W., Woods, E. R. (1996). Comparison of two violence prevention curricula for middle school adolescents. *Journal of Adolescent Health, 19*, 111-117.
- Ellis, A. K., & Fouts, J. T. (1994). Research on school restructuring. Princeton, NJ: The Library of Innovations.
- Farrell, A. D., Meyer, A. L., & White, K. S. (2001). Evaluation of Responding in Peaceful and Positive Ways (RIPP): A school-based prevention program for reducing violence among urban adolescents. *Journal of Clinical Child Psychology, 30*(4), 451-463.
- Ginsberg, R., & Berry, B. (1998). The capability for enhancing accountability. Educational Policy, 12, 48-66.
- Hall G. E., & Hord S. M. (2001). *Implementing Change*. Boston: Allyn and Bacon.
- Hausman, A., Pierce, G., & Briggs, L. (1996). Evaluation of comprehensive violence prevention education: Effects on student behavior.
- Hausman, A. J., Spivak, H., & Prothrow-Stith, D. (1995). Evaluation of a community-based youth violence prevention project. *Journal of Adolescent Health, 17*, 353-359.
- Kirst M. (1990). Accountability: Implications for state and local policymakers. Washington, DC: US Department of Education.
- Levin, H. (1974). A conceptual framework for accountability in education. School Review, 82, 363-391.
- Mytton, J. A., DiGuseppi, C., Gough, D. A., Taylor, R. S., & Logan, S. (2002). School-based violence prevention programs. *Archives of Pediatric and Adolescent Medicine, 156*, 752-762.

- Munoz, M. A. (Fall 2002). Accountability in education: Equity as learning for all. Reality Test. American Educational Research Association, Division H School Evaluation and Program Development. Available on-line at www.aera.net/divisions/h
- Murphy, J. (1988). Equity as student opportunity to learn. *Theory Into Practice*, 27, 145-151.
- National Commission on Excellence in Education (1983). A nation at risk. Washington, DC: Author.
- Nelson, J. L., Palonsky, S. B., & Carlson, K. (1990). *Critical issues in education*. New York: McGraw-Hill.
- Oprinas, P., Parcel, G. S., McAlister, A., & Frankowski, R. (1995). Violence prevention in middle schools: A pilot evaluation. *Journal of Adolescent Health*, 17, 360-371.
- Pinkney, H. B. (1980). An American dilemma: Financing public education. *NASSP Bulletin*, 64, 68-73.
- Rich, J. M. (1985). *Innovations in education: Reformers and their critics*. Boston: Allyn and Bacon.
- Rossi, P. H., Freeman, H. E., & Lipsey, M. W. (1999). *Evaluation: A systematic approach* (6th ed.). Thousand Oaks, CA: Sage.
- Tabachnik, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York, NY: HarperCollins Publishers, Inc.
- Twemlow, S. W., Fonagy, P., Sacco, F. C., Gies, M. L., Evans, R., & Ewbank, R. (2001). Creating a peaceful school learning environment: A controlled study of an elementary school intervention to reduce violence. *American Journal of Psychiatry*, 158(5), 808-810.
- Webster, D. (1991). The unconvincing case for school-based conflict resolution. *Health Affairs*, 12, 126-141.

Wohlstetter, P. (1991). Accountability mechanisms for state education reform: Some organizational alternatives. *Educational Evaluation and Policy Analysis*, 13, 31-48.

APPENDIX A

Table 3. Short-term results of intervention impacts on non-cognitive and cognitive outcomes for Cohorts 1 and 2. GPA not reported for students at the elementary level (Second Step, BB/BS, and PMH).

Cohort 1 Outcomes	SMART Moves (n=1814)		Control (n=1708)		F	Eta
	Spring 00	Spring 01	Spring 00	Spring 01		
Unexcused absence	4.88	5.08	3.68	3.55	2.39	0.03
Unexcused tardiness	4.65	7.40	3.28	3.88	57.92***	0.13
ISAPs	.35	.84	.50	1.28	9.36**	0.05
Suspensions	.18	.32	.12	.16	16.80***	0.07
Days suspended	.51	1.05	.39	.52	19.75***	0.08
GPA	2.03	2.02	2.10	2.03	2.32	0.03
Cohort 2 Outcomes	SMART Moves (n=2416)		Control (n=2329)		F	Eta
	Spring 01	Spring 02	Spring 01	Spring 02		
Unexcused absence	4.97	4.39	2.92	2.92	9.66**	0.04
Unexcused tardiness	5.96	4.88	3.41	3.42	24.92***	0.07
ISAPs	1.2	1.43	0.78	1.13	1.05	0
Suspensions	0.32	0.21	0.08	0.14	65.23***	0.12
Days suspended	1.06	0.79	0.27	0.47	35***	0.08
GPA	1.92	1.85	2.2	2.2	5.16*	0.04
Cohort 1 Outcomes	Second Step (n=922)		Control (n=918)		F	Eta
	Spring 00	Spring 01	Spring 00	Spring 01		
Unexcused absence	5.21	3.05	2.98	2.25	33.70***	0.13
Unexcused tardiness	4.78	4.47	3.12	3.04	0.53	0
ISAPs	0.001	0	0	0	1.07	0.03
Suspensions	0.02	0.02	0.01	0.03	0.771	0
Days suspended	0.008	0.007	0.004	0.01	0.832	0
Cohort 2 Outcomes	Second Step (n=1225)		Control (n=1200)		F	Eta
	Spring 01	Spring 02	Spring 01	Spring 02		
Unexcused absence	3.09	2.28	2.3	1.97	7.21**	0.05
Unexcused tardiness	3.91	3.4	2.43	2.63	9.13**	0.06
ISAPs	0	0	0	0	0***	0
Suspensions	0.01	0.01	0.01	0.02	0.17	0
Days suspended	0.02	0.04	0.03	0.06	0.23	0
Cohort 1 Outcomes	LEEP (n=119)		Control (n=102)		F	Eta
	Spring 00	Spring 01	Spring 00	Spring 01		
Unexcused absence	8.47	9.42	5.33	6.13	0.011	0
Unexcused tardiness	10.56	8.69	3.68	6.12	6.32*	0.17
ISAPs	1.34	1.29	2.81	2.05	1.03	0.07
Suspensions	0.57	0.65	0.21	0.32	0.005	0
Days suspended	0.17	0.19	0.05	0.09	0.005	0
GPA	1.65	1.68	1.71	1.94	1.35	0.08
Cohort 2 Outcomes	LEEP (n=100)		Control (n=97)		F	Eta
	Spring 01	Spring 02	Spring 01	Spring 02		
Unexcused absence	9.34	5.83	6.88	5.95	2.79	0.12
Unexcused tardiness	8.18	8.45	6.21	4.46	1.36	0.08
ISAPs	1.12	0.51	2.05	1.57	0.08	0
Suspensions	0.2	0.09	0.2	0.13	0.18	0.03
Days suspended	0.63	0.36	0.92	0.34	0.61	0.05
GPA	1.73	1.78	1.75	1.9	0.28	0.04

Table continues...

Table 3 continued. Short-term results of intervention impacts on non-cognitive and cognitive outcomes for Cohorts 1 and 2. GPA not reported for students at the elementary level (Second Step, BB/BS, and PMH).

Cohort 1 Outcomes	FFT (n=75)		Control (n=69)		F	Eta
	Spring 00	Spring 01	Spring 00	Spring 01		
Unexcused absence	5.01	7.75	4	3.47	6.59*	0.21
Unexcused tardiness	4.43	8.19	2.74	3.03	5.39*	0.19
ISAPs	1.05	0.91	0.82	1.3	1.08	0.09
Suspensions	1.58	2.66	0.64	0.31	4.51*	0.18
Days suspended	0.5	0.68	0.19	0.09	2.15	0.12
GPA	1.53	1.82	1.86	1.92	0.735	0.11
Cohort 2 Outcomes	FFT (n=25)		Control (n=24)		F	Eta
	Spring 01	Spring 02	Spring 01	Spring 02		
Unexcused absence	5.15	4.98	1.92	3.6	0.64	0.12
Unexcused tardiness	5.61	5.52	5.24	4.08	0.35	0.09
ISAPs	1.08	1.06	1.05	0.98	0.01	0
Suspensions	0.54	0.53	0.14	0	0.23	0.07
Days suspended	2.25	1.66	0.37	0.02	0.03	0.03
GPA	1.87	1.66	2.44	2.05	0.24	0.09
Cohort 1 Outcomes	PMH (n=233)		Control (n=233)		F	Eta
	Spring 00	Spring 01	Spring 00	Spring 01		
Unexcused absence	4.47	3.33	3	1.96	0.047	0
Unexcused tardiness	4.81	4.89	3.35	3	0.453	0.03
ISAPs	0	0.03	0	0.01	0.011	0
Suspensions	0	0.04	0	0.03	0.015	0
Days suspended	0	0.01	0	0.02	0.184	0
Cohort 2 Outcomes	PMH (n=450)		Control (n=443)		F	Eta
	Spring 01	Spring 02	Spring 01	Spring 02		
Unexcused absence	3.42	2.75	2.43	1.92	0.4	0
Unexcused tardiness	4.11	3.51	2.45	2.92	7.81**	0.09
ISAPs	0	0	0	0	0.49	0
Suspensions	0.01	0.01	0.01	0.01	0.5	0.03
Days suspended	0.02	0.01	0.03	0.04	0.2	0
Cohort 1 Outcomes	BB/BS (n=152)		Control (n=151)		F	Eta
	Spring 00	Spring 01	Spring 00	Spring 01		
Unexcused absence	3.81	2.72	2.84	2.05	0.419	0.03
Unexcused tardiness	5.9	4.12	2.33	2.64	9.27**	0.17
ISAPs	0	0.05	0	0.02	0.443	0.03
Suspensions	0.03	0.09	0	0.07	0.028	0
Days suspended	0.02	0.04	0	0.02	0	0
Cohort 2 Outcomes	BB/BS (n=180)		Control (n=179)		F	Eta
	Spring 01	Spring 02	Spring 01	Spring 02		
Unexcused absence	2.7	2.8	2.35	2.02	1.54	0.06
Unexcused tardiness	4.02	3.37	2.71	2.47	0.42	0.03
ISAPs	0.05	0.3	0.2	0.15	2.31	0.08
Suspensions	0.06	0.11	0.08	0.05	2.43	0.08
Days suspended	0.12	0.38	0.17	0.26	0.85	0.04

Table continues...

Table 3 continued. Short-term results of intervention impacts on non-cognitive and cognitive outcomes for Cohorts 1 and 2. GPA not reported for students at the elementary level (Second Step, BB/BS, and PMH).

Cohort 1 Outcomes	MST (n=57)		Control (n=52)		F	Eta
	Spring 00	Spring 01	Spring 00	Spring 01		
Unexcused absence	11.36	9.35	6.11	7.74	2.25	0.15
Unexcused tardiness	7.98	5.6	4.67	8.21	5.82*	0.23
ISAPs	1.88	1.12	2.76	3.26	1.26	0.11
Suspensions	1.9	1.83	0.47	0.76	0.111	0.03
Days suspended	0.53	0.44	0.17	0.26	0.442	0.06
GPA	0.8	1.01	1.58	1.65	0.146	0.04

Cohort 2 Outcomes	MST (n=36)		Control (n=34)		F	Eta
	Spring 01	Spring 02	Spring 01	Spring 02		
Unexcused absence	12.07	8.14	5.84	4.44	0.72	0.1
Unexcused tardiness	9.38	12.83	4.25	4.62	1.13	0.13
ISAPs	2.54	2.95	2.05	3.67	0.53	0.09
Suspensions	0.57	0.46	0.16	0.07	0.01	0
Days suspended	2.21	1.67	0.87	0.26	0	0
GPA	1.04	1.18	1.91	1.94	0.18	0.06

NOTE: LEEP = Leadership through Experiential Education Program; FFT = Functional Family Therapy; PMH = Primary Mental Health; BB/BS=Big Brothers/Big Sisters; MST = Multi-Systemic Therapy. *** $p < .001$, ** $p < .01$, * $p < .05$. Because of small sample sizes, there was insufficient power to detect significant changes.

Table 4. Long-term results of intervention impacts on non-cognitive and cognitive outcomes for Cohorts 1 and 2. GPA not reported for students at the elementary level (Second Step, BB/BS, and PMH).

Cohort 1 Outcomes	SMART Moves (n=1814)		Control (n=1708)		F	Eta
	Spring 00	Spring 02	Spring 00	Spring 02		
Unexcused absence	4.88	9.43	3.68	6.99	10.51***	0.05
Unexcused tardiness	4.65	10.27	3.28	8.02	3.44	0.03
ISAPs	0.35	2.42	0.5	3.13	5.47*	0.04
Suspensions	0.51	1.44	0.39	1.45	0.58	0
Days suspended	0.18	0.4	0.12	0.38	1.05	0
GPA	2.06	1.93	2.1	2.03	1.49	0.03

Cohort 2 Outcomes	SMART Moves (n=2416)		Control (n=2329)		F	Eta
	Spring 01	Spring 03	Spring 01	Spring 03		
Unexcused absence	4.97	4.86	2.92	3.73	16.83***	0.06
Unexcused tardiness	5.96	5.17	3.41	4.32	45.86***	0.1
ISAPs	1.2	1.79	0.78	1.44	0.22	0
Suspensions	0.32	0.22	0.08	0.17	67.06***	0.12
Days suspended	1.06	0.72	0.27	0.57	55.12***	0.1
GPA	1.7	0.61	2.17	2.04	33.8***	0.39

Cohort 1 Outcomes	Second Step (n=922)		Control (n=918)		F	Eta
	Spring 00	Spring 02	Spring 00	Spring 02		
Unexcused absence	5.21	4.2	3	3.83	35.96***	0.14
Unexcused tardiness	4.78	6.83	3.14	5.75	1.21	0.03
ISAPs	0	0	0	0	1.07	0.03
Suspensions	0.02	0.07	0.01	0.1	0.315	0
Days suspended	0	0.03	0	0.02	0.018	0

Table continues...

Table 4 continued. Long-term results of intervention impacts on non-cognitive and cognitive outcomes for Cohorts 1 and 2. GPA not reported for students at the elementary level (Second Step, BB/BS, and PMH).

Cohort 2 Outcomes	Second Step (n=1225)		Control (n=1200)			
	Spring 01	Spring 03	Spring 01	Spring 03		
Unexcused absence	3.09	2.65	2.3	2.23	4.4*	0.04
Unexcused tardiness	3.91	3.27	2.43	2.67	12.16***	0.07
ISAPs	0	0	0	0.02	3.89*	0.04
Suspensions	0.01	0.01	0.01	0.01	0.04	0
Days suspended	0.02	0.03	0.03	0.05	0.09	0
Cohort 1 Outcomes	LEEP (n=119)		Control (n=102)		F	Eta
	Spring 00	Spring 02	Spring 00	Spring 02		
Unexcused absence	8.65	10.25	5.14	11.53	3.32	0.13
Unexcused tardiness	10.82	14.04	3.64	11.01	1.36	0.08
ISAPs	1.41	0.2	2.71	3.99	4.94*	0.15
Suspensions	0.58	0.67	0.22	0.46	0.13	0.03
Days suspended	0.16	0.16	0.06	0.23	1.98	0.09
GPA	1.66	1.68	1.69	2.11	2.55	0.15
Cohort 2 Outcomes	LEEP (n=100)		Control (n=97)			
	Spring 01	Spring 03	Spring 01	Spring 03		
Unexcused absence	9.34	4.17	6.88	5.96	6.15*	0.18
Unexcused tardiness	8.18	6.49	6.21	5.8	0.47	0.04
ISAPs	1.12	0.34	2.05	0.79	1.28	0.08
Suspensions	0.2	0.07	0.2	0.17	1.18	0.08
Days suspended	0.63	0.33	0.92	0.56	0.03	0
GPA	1.26	2.11	1.75	2.06	0.33	0.08
Cohort 1 Outcomes	FFT (n=75)		Control (n=69)		F	Eta
	Spring 00	Spring 02	Spring 00	Spring 02		
Unexcused absence	4.82	10.05	4.68	10.77	0.059	0.03
Unexcused tardiness	5.06	8.74	3.3	8.78	0.476	0.07
ISAPs	0.99	1.71	0.78	4.64	3.75*	0.2
Suspensions	1.7	2.37	0.85	1.69	0.023	0
Days suspended	0.48	0.77	0.22	0.46	0.017	0
GPA	1.34	1.53	1.92	2.27	0.254	0.06
Cohort 2 Outcomes	FFT (n=25)		Control (n=24)			
	Spring 01	Spring 03	Spring 01	Spring 03		
Unexcused absence	5.15	4.98	1.92	3.02	0.34	0.08
Unexcused tardiness	5.61	4.5	5.24	2.85	0.31	0.08
ISAPs	1.08	1.44	1.05	1.58	0.02	0
Suspensions	0.54	0.17	0.14	0.24	2.21	0.21
Days suspended	2.25	0.66	0.37	0.4	1.65	0.19
GPA	1.98	1	2.3	2.21	2.4	0.33
Cohort 1 Outcomes	PMH (n=233)		Control (n=233)		F	Eta
	Spring 00	Spring 02	Spring 00	Spring 02		
Unexcused absence	4.04	5.82	2.82	3.65	1.27	0.07
Unexcused tardiness	4.26	6.21	2.95	4.87	0	0
ISAPs	0	0	0	0	0	0
Suspensions	0	0.08	0	0.01	1.13	0.06
Days suspended	0	0.05	0	0.01	0.525	0.04

Table continues...

Table 4 continued. Long-term results of intervention impacts on non-cognitive and cognitive outcomes for Cohorts 1 and 2. GPA not reported for students at the elementary level (Second Step, BB/BS, and PMH).

Cohort 2 Outcomes	PMH (n=450)		Control (n=443)		F	Eta
	Spring 01	Spring 03	Spring 01	Spring 03		
Unexcused absence	3.42	2.96	2.43	2.27	0.97	0.03
Unexcused tardiness	4.11	4.03	2.45	3.1	2.82	0.05
ISAPs	0	0	0	0	0.76	0.03
Suspensions	0.01	0.01	0.01	0.01	0.34	0
Days suspended	0.02	0.04	0.03	0.08	0.27	0
Cohort 1 Outcomes	BB/BS (n=152)		Control (n=151)		F	Eta
	Spring 00	Spring 02	Spring 00	Spring 02		
Unexcused absence	3.67	5.66	2.94	3.59	3.53	0.11
Unexcused tardiness	6.01	7.52	2.43	5.17	0.779	0.05
ISAPs	0	0.25	0	0.75	3.79	0.11
Suspensions	0.04	0.84	0	0.46	0.853	0.05
Days suspended	0.02	0.23	0	0.15	0.301	0.03
Cohort 2 Outcomes	BB/BS (n=180)		Control (n=179)		F	Eta
	Spring 01	Spring 03	Spring 01	Spring 03		
Unexcused absence	2.7	3	2.35	2.55	0.07	0
Unexcused tardiness	4.02	3.63	2.71	3.46	2.22	0.08
ISAPs	0.05	0.29	0.2	0.6	0.95	0.05
Suspensions	0.06	0.18	0.08	0.11	1.74	0.07
Days suspended	0.12	0.65	0.17	0.39	1.8	0.07
Cohort 1 Outcomes	MST (n=57)		Control (n=52)		F	Eta
	Spring 00	Spring 02	Spring 00	Spring 02		
Unexcused absence	11.95	7.36	5.92	13.37	12.41***	0.33
Unexcused tardiness	8.36	11.81	4.58	11.24	0.442	0.06
ISAPs	2.01	2.42	2.7	2.6	0.1	0.03
Suspensions	1.9	0.62	0.47	1.45	4.67*	0.21
Days suspended	0.53	0.13	0.17	0.52	6.92**	0.25
GPA	1.11	1.12	1.68	1.61	0.029	0.03
Cohort 2 Outcomes	MST (n=36)		Control (n=34)		F	Eta
	Spring 01	Spring 03	Spring 01	Spring 03		
Unexcused absence	12.07	4.59	5.84	3.66	2.63	0.19
Unexcused tardiness	9.38	8.11	4.25	5.35	0.59	0.09
ISAPs	2.54	0.8	2.05	2.92	3.18	0.21
Suspensions	0.57	0.08	0.16	0.09	3.52	0.22
Days suspended	2.21	0.33	0.87	0.56	2.03	0.17
GPA	2.86	2.2	2.48	2.41	0.22	0.13

NOTE: LEEP = Leadership through Experiential Education Program; FFT = Functional Family Therapy; PMH = Primary Mental Health; BB/BS=Big Brothers/Big Sisters; MST = Multi-Systemic Therapy. *** $p < .001$, ** $p < .01$, * $p < .05$. Because of small sample sizes, there was insufficient power to detect significant changes.

Table 5. The relationship between dosage and changes in short term and long term outcomes for students in Cohort 2.

Intervention	Short Term		Long Term	
	B	Beta	B	Beta
<u>Big Brothers/Big Sisters</u>				
ISAPs	—		0.5	
Dosage	—	—	-0.01	-0.15*
<u>LEEP Mentoring</u>				
Absences	-12.03		-15.32	
Dosage	0.12	0.4***	0.14	0.42***
Tardiness	-10.04		-13.24	
Dosage	0.14	0.33***	0.16	0.37***
Suspensions	0.24		—	
Dosage	-0.01	-0.26**	—	—
Days Suspended	1.23		—	
Dosage	-0.02	-0.29**	—	—
<u>Multi-Systemic Therapy</u>				
Suspensions	-1.17		—	
Dosage	0.03	0.35*	—	—
Days Suspended	-5.43		—	
Dosage	0.15	0.41*	—	—
<u>Second Step</u>				
Tardiness	-1.06		-1.46	
Dosage	0.02	0.06*	0.02	0.08**

NOTE: B = change in outcomes (number of days); positive beta denotes an increase in dosage. Dosage is defined as the number of weeks the student received services in the intervention. *** $p < .001$, ** $p < .01$, * $p < .05$. Short term change = Spring 2001 – Spring 2002. Long term change = Spring 2001 – Spring 2003.



U.S. Department of Education
 Office of Educational Research and Improvement
 (OERI)
 National Library of Education (NLE)
 Educational Resources Information Center (ERIC)



Reproduction Release

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: <i>Pushing the Boundaries of Education</i>	
Author(s): <i>William Neace et al</i>	
Corporate Source: <i>JCPS / PIRE</i>	Publication Date: <i>2003</i>

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign in the indicated space following.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY <i>SAMPLE</i> TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY <i>SAMPLE</i> TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY <i>SAMPLE</i> TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
Level 1	Level 2A	Level 2B
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g. electronic) and paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only
Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.		

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche, or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: Marco Minor	Printed Name/Position/Title: Marco A. Minor Evaluation Specialist
Organization/Address: 3332 Newburg Rd Louisville, KY 40218	Telephone: (502) 485-6348 Fax: (502) 485-6255
	E-mail Address: Date: 09/28/2003

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse: